

Shared Control of Human and Robot by Approximate Dynamic Programming

Yanan Li, Keng Peng Tee, Dilip Limbu, and Shuzhi Sam Ge

Abstract— This paper aims at proposing a general framework of human-robot shared control for a natural and effective interface. A typical human-robot collaboration scenario is investigated, and a framework of shared control is developed based on finding the solution to an optimization problem. Human dynamics are taken into account in the analysis of the coupled human-robot system, and objectives of both human and robot are considered. Approximate dynamic programming is employed to solve the optimization problem in the presence of unknown human and robot dynamics. The validity of the proposed method is verified through simulation studies.

I. INTRODUCTION

Despite the advent of robotics in the last decades, the development of fully autonomous robots that fulfil operational requirements under real-world working conditions is still very challenging. The intervention of human beings in a task is necessary especially when the environment is unstructured and new to robots. Many researchers have seen the need for shared control of human and robot in social applications, industrial settings, and space explorations, among others.

In the early literature of human-robot shared control, a leader-follower model is usually adopted wherein the robot is assigned a follower's role [1]. Indeed, human beings' superb capabilities of situation awareness and decision making naturally predispose them to take on a leader's role. Under this model, variable impedance control is developed by estimating impedance parameters of the human arm and adjusting impedance parameters of the robot arm [2], [3]. Recently, many researchers have recognized the importance of a framework beyond the leader-follower model [4]. Such a framework is expected to assign adjustable roles to human and robot, such that human effort can be reduced and task performance improved.

The initial effort in this direction is the prediction of human intention, based on which the robot switches/adjusts its role when necessary. In [5], the knowledge of motion characteristics of the human arm is used to generate a point-to-point movement for the robot to collaborate with the human partner. In [6], intention recognition based on haptic

data has been investigated for automatic activation of assistance. In [7], human's motion intention is predicted based on a model acquired using probabilistic learning methods with the technique of learning by demonstration. In [8], human's motion intention is represented by the change of the interaction force. In [9], the intentional walking direction of the users of a crane robot is estimated using a Kalman filter. In our previous work [10], the online estimation of human's motion intention is achieved based on the dynamic model of the human arm.

In aforementioned works, a robot is expected to provide assistance to a human being while it does not have its own objective. In many cases, however, a robot should also have its own objective, e.g., a robot follows a predefined trajectory while the human being intervenes by driving it through several points-of-interest. To date, very little effort has been made to reach this target. In [11], based on the objective of minimizing the human effort, a role adaptation strategy is developed to switch between model-based and model-free predictions in the case of partially known tasks. In [12], [13], the terminal constraint receding horizon control is developed for shared control of human and robot, such that the common goal can be achieved subsequently. A homotopy switching model is developed in [14] for dyad haptic interaction in physical collaborative tasks. In [15], a generic framework is proposed to analyze interactive motor behaviors, where system dynamics are required to be known. In this paper, we aim to develop a framework of human-robot shared control with both objectives of human and robot taken into account. Under this framework, approximate dynamic programming will be employed to solve the optimal problem subject to unknown system dynamics [16], [17].

The rest of this paper is organized as follows. In Section II, a typical human-robot collaboration scenario is introduced, and the problem of shared control of human and robot is formulated. In Section III, the system under study is described in the state-space form, shared control is formulated as an optimization problem, and approximate dynamic programming is employed to cope with the problem of unknown human and robot dynamics. Results of case studies are shown in Section IV, and concluding remarks are given in Section V.

II. PROBLEM FORMULATION

A. System Description

In this paper, we consider a scenario where a human arm physically interacts with a robot arm. There is a force/torque sensor at the interaction point so the measurement of the

This work is supported by Grant No. 1225100001 from the Science and Engineering Research Council (SERC), A*STAR, Singapore.

Y. Li, K. P. Tee, and D. Limbu are with the Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore 138632. {liy,kptee,dklimbu}@i2r.a-star.edu.sg

Shuzhi Sam Ge is with the Department of Electrical and Computer Engineering and the Social Robotics Laboratory, Interactive Digital Media Institute, National University of Singapore, Singapore 117576. samge@nus.edu.sg

interaction force (including force and torque) between human and robot is available. This scenario can be found in applications such as robotic welding and object transporting [18].

The kinematic relationship of the robot arm between the joint space and the Cartesian space (interaction point) can be described as

$$x(t) = \phi(q) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is position/oritation in the Cartesian space, and $q \in \mathbb{R}^n$ is coordinate in the joint space, respectively. Differentiating Equ. (1) with respect to time leads to

$$\dot{x}(t) = J(q)\dot{q} \quad (2)$$

where $J(q) \in \mathbb{R}^{n \times n}$ is the Jacobian matrix which is assumed to be known and nonsingular in a finite workspace. Further differentiating Equ. (2) with respect to time results in

$$\ddot{x}(t) = \dot{J}(q)\dot{q} + J(q)\ddot{q} \quad (3)$$

The dynamic model of the robot arm in the joint space is given by

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau + J^T(q)f \quad (4)$$

where $M(q) \in \mathbb{R}^{n \times n}$ is the symmetric bounded positive definite inertia matrix, $C(q, \dot{q}) \in \mathbb{R}^n$ is the Coriolis and Centrifugal force, $G(q) \in \mathbb{R}^n$ is the gravitational force, $\tau \in \mathbb{R}^n$ is the control input, and $f \in \mathbb{R}^n$ is the interaction force between human and robot which is in the Cartesian space. Substituting the kinematic constraints (1), (2), and (3) into the dynamic model (4), we obtain the dynamic model of the robot arm in the Cartesian space, as below

$$M_R(q)\ddot{x} + C_R(q, \dot{q})\dot{x} + G_R(q) = u + f \quad (5)$$

where

$$\begin{aligned} M_R(q) &= J^{-T}(q)M(q)J^{-1}(q) \\ C_R(q, \dot{q}) &= J^{-T}(q)(C(q, \dot{q}) - M(q)J^{-1}(q)\dot{J}(q))J^{-1}(q) \\ G_R(q) &= J^{-T}(q)G(q), \quad u = J^{-T}(q)\tau \end{aligned} \quad (6)$$

Since only the control input u is designable, we treat the human arm as a system that

- applies a force f to control the states of the robot arm, and simultaneously
- changes its own states according to the states of the robot arm and the interaction force f .

To describe the above interpretation in a rigorous manner, we employ the following dynamic model of the human arm under the equilibrium point control hypothesis [19], [20], [21]:

$$M_H\ddot{x} + C_H\dot{x} + K_H(x - x_H) = -f \quad (7)$$

where M_H , C_H , and K_H are inertia, damping, and stiffness matrices of the human arm, respectively, and x_H is the equilibrium position as planned in human's central nervous system.

Remark 1: This model suggests that the interaction force f drives the human arm away from the equilibrium position,

but it does not reflect how human controls his/her arm to make it behave in this way. The topic of the so-called human motor control is out of scope of this paper, and interested readers may refer to references [22], [23], [24]. Also note that in the scenario under study in this paper, the interaction between human and robot is kinesthetic, so the position x of the human arm and robot arm is identical.

In [25], [3], it has been argued that the inertia M_H in Equ. (7) is negligible, while the damping C_H and stiffness K_H dominate the dynamics of the human arm. Following these arguments, we consider the following simplified model to describe the dynamics of the human arm:

$$C_H\dot{x} + K_H(x - x_H) = -f \quad (8)$$

Note that C_H and K_H are time-varying because human may modulate the damping and stiffness of his/her arm in different stages of interaction. In particular, the damping matrix C_H is due to damping and Coriolis effects, so it is a function of x and \dot{x} . Referring to the stiffness ellipse in [20], [21], the stiffness matrix K_H is position- and velocity-dependent, so it is also a function of x and \dot{x} . In summary, we have the following property:

Property 1: C_H and K_H are functions of x and \dot{x} .

Lemma 1: Given a vector $\xi \in \mathbb{R}^n$, we have

$$\begin{aligned} \dot{C}_H\xi &= C_{H1}(x, \xi)\dot{x} + C_{H2}(\dot{x}, \xi)\ddot{x} \\ \dot{K}_H\xi &= K_{H1}(x, \xi)\dot{x} + K_{H2}(\dot{x}, \xi)\ddot{x} \end{aligned} \quad (9)$$

where $C_{H1}(x, \xi)$, $C_{H2}(\dot{x}, \xi)$, $K_{H1}(x, \xi)$, and $K_{H2}(\dot{x}, \xi)$ are all $n \times n$ matrices, of which the particular forms are given in the following proof.

Proof: Denote $\rho = \dot{C}_H\xi$, the vector composed by elements at the i -th row of C_H as C_i , and the element of C_H at the i -th row and j -th column as c_{ij} . According to Property 1, we have $\dot{c}_{ij} = \frac{\partial c_{ij}}{\partial x}\dot{x} + \frac{\partial c_{ij}}{\partial \dot{x}}\ddot{x}$.

Consider the i -th element of ρ as below

$$\begin{aligned} \rho_i &= \sum_{j=1}^n \dot{c}_{ij}\xi_j \\ &= \sum_{j=1}^n \left[\left(\frac{\partial c_{ij}}{\partial x}\dot{x} + \frac{\partial c_{ij}}{\partial \dot{x}}\ddot{x} \right) \xi_j \right] \\ &= \sum_{j=1}^n \sum_{k=1}^n \left[\left(\frac{\partial c_{ij}}{\partial x_k}\dot{x}_k + \frac{\partial c_{ij}}{\partial \dot{x}_k}\ddot{x}_k \right) \xi_j \right] \\ &= \sum_{k=1}^n \sum_{j=1}^n \left[\left(\frac{\partial c_{ij}}{\partial x_k}\dot{x}_k + \frac{\partial c_{ij}}{\partial \dot{x}_k}\ddot{x}_k \right) \xi_j \right] \\ &= \sum_{k=1}^n \left(\frac{\partial C_i}{\partial x_k}\xi\dot{x}_k + \frac{\partial C_i}{\partial \dot{x}_k}\xi\ddot{x}_k \right) \\ &= \frac{\partial C_i}{\partial x^T}\xi\dot{x} + \frac{\partial C_i}{\partial \dot{x}^T}\xi\ddot{x} \end{aligned} \quad (10)$$

From the above equation, we find that the i -th rows of $C_{H1}(x, \xi)$ and $C_{H2}(\dot{x}, \xi)$ are $\frac{\partial C_i}{\partial x^T}\xi$ and $\frac{\partial C_i}{\partial \dot{x}^T}\xi$, respectively. Therefore, $\dot{C}_H\xi = C_{H1}(x, \xi)\dot{x} + C_{H2}(\dot{x}, \xi)\ddot{x}$. Similarly, it can be shown that $\dot{K}_H\xi = K_{H1}(x, \xi)\dot{x} + K_{H2}(\dot{x}, \xi)\ddot{x}$, of which the details are omitted.

B. Control Objective

As discussed in Introduction, the control framework under development is beyond the simple leader-follower model.

In particular, we consider that both human and robot have their own objectives which may be coherent or conflicting. The overall control objective should be the trade-off of both individual objectives, described by the following cost function:

$$\Gamma(t) = \int_t^\infty c(v)dv \quad (11)$$

where $c(t)$ is the instant cost function defined as

$$c(t) = (x - x_d)^T Q_1 (x - x_d) + \dot{x}^T Q_2 \dot{x} + f^T Q_3 f + u^T R u \quad (12)$$

In the above definition, x_d is the desired trajectory of the robot arm, $Q_1 \in \mathbb{R}^{n \times n} \geq 0$, $Q_2 \in \mathbb{R}^{n \times n} \geq 0$, and $Q_3 \in \mathbb{R}^{n \times n} \geq 0$ are the weights of position tracking, velocity, and interaction force, respectively, and $R \in \mathbb{R}^{n \times n} > 0$ is the weight of the control input of the robot arm.

Remark 2: The last term of the cost function Γ penalizes the control input of the robot arm. The first two terms penalize the error between the actual position (velocity) and the desired position (zero velocity) of the robot arm, and it indicates that the objective of the robot arm is “respected”. Similarly, there should have been a term to penalize the error between the actual position and the desired position of the human arm, i.e., $(x - x_H)$. However, x_H is unmeasurable, so it is replaced by the interaction force f . The rationale of the replacement can be understood from the dynamic model (8), which indicates that f is a measure of conflict between human intention and actual position.

III. SHARED CONTROL

This section is dedicated to introduce the details of designing the control input of the robot arm u , such that the cost function (11) is minimized. Two main steps will be presented: the transformation of the system dynamics to a state-space form will be conducted, and an optimal tracking problem in absence of unknown nonlinear dynamics will be solved by approximate dynamic programming.

A. State-Space Form

Taking the derivative of the human dynamics (8) with respect to time leads to

$$\dot{C}_H \dot{x} + C_H \ddot{x} + \dot{K}_H (x - x_H) + K_H \dot{x} = -\dot{f} \quad (13)$$

Rearranging Equ. (13) and considering Equ. (8) and Property 1, we have

$$\begin{aligned} \dot{f} &= -\dot{K}_H (x - x_H) - K_H \dot{x} - \dot{C}_H \dot{x} - C_H \ddot{x} \\ &= -K_{H1} (x, x - x_H) \dot{x} - K_{H2} (\dot{x}, x - x_H) \ddot{x} - K_H \dot{x} \\ &\quad - C_{H1} (x, \dot{x}) \dot{x} - C_{H2} (\dot{x}) \ddot{x} - C_H \ddot{x} \end{aligned} \quad (14)$$

For convenience, the above equation is rewritten as

$$\dot{f} = -K_{H1} \dot{x} - K_{H2} \ddot{x} - K_H \dot{x} - C_{H1} \dot{x} - C_{H2} \ddot{x} - C_H \ddot{x} \quad (15)$$

Rearranging Equ. (5), we have

$$\ddot{x} = -M_R^{-1}(q)C_R(q, \dot{q})\dot{x} - M_R^{-1}(q)G_R(q) + M_R^{-1}(q)f + M_R^{-1}(q)u \quad (16)$$

Choose the following three states $z_1 = x$, $z_2 = \dot{x}$, and $z_3 = f$ to form the system state $z = [z_1, z_2, z_3]^T$. Considering Eqs. (15) and (16), the system dynamics can be described in the following state-space form

$$\dot{z} = A(z)z + B(z)u + D(z) \quad (17)$$

where

$$\begin{aligned} A(z) &= L^{-1} \begin{bmatrix} \mathbf{0}_n & I_n & \mathbf{0}_n \\ \mathbf{0}_n & -M_R^{-1}(q)C_R(q, \dot{q}) & M_R^{-1}(q) \\ \mathbf{0}_n & -K_{H1} - K_H - C_{H1} & \mathbf{0}_n \end{bmatrix} \\ B(z) &= L^{-1} \begin{bmatrix} \mathbf{0}_n \\ M_R^{-1}(q) \\ \mathbf{0}_n \end{bmatrix} \\ D(z) &= L^{-1} \begin{bmatrix} \mathbf{0}_n \\ -M_R^{-1}(q)G_R(q) \\ \mathbf{0}_n \end{bmatrix} \end{aligned} \quad (18)$$

with

$$L = \begin{bmatrix} I_n & \mathbf{0}_n & \mathbf{0}_n \\ \mathbf{0}_n & I_n & \mathbf{0}_n \\ \mathbf{0}_n & C_{H2} + C_H + K_{H2} & I_n \end{bmatrix} \quad (19)$$

$\mathbf{0}_n$ and I_n denote $n \times n$ zero and identity matrices, respectively. Considering $q = \phi^{-1}(x)$, $\dot{q} = J^{-1}\dot{x}$, $x - x_H = K_H^{-1}(f + C_H \dot{x})$, and Property 1, the argument of A , B , and D is z . Besides, L^{-1} always exists.

Remark 3: The interaction force f is deemed as a state in the coupled system since it is a measurable variable from the human side. In this way, the effect of the states of human and robot on each other, i.e., concurrent adaptation, is described in an analytic way. This is the feature of the proposed framework of shared control, which is different from existing methods such as [5], [6], [7], [8], [9], [10].

The above optimal tracking problem needs to be transformed to a regulation problem such that available optimal control can be employed. In particular, we consider that the desired trajectory of the robot arm x_d is generated from the following system

$$\begin{cases} \dot{w} = Ew \\ x_d = Fw \end{cases} \quad (20)$$

where $w \in \mathbb{R}^l$ is an auxiliary state, and $E \in \mathbb{R}^{l \times l}$ and $F \in \mathbb{R}^{n \times l}$ are two known matrices and (E, F) is observable.

Considering the augmented state $\bar{z} = [z^T \ w^T]^T$, we have the following augmented system

$$\dot{\bar{z}} = \bar{A}(\bar{z})\bar{z} + \bar{B}(\bar{z})u + \bar{D}(\bar{z}) \quad (21)$$

where

$$\begin{aligned} \bar{A}(\bar{z}) &= \begin{bmatrix} A(z) & \mathbf{0}_{3n \times l} \\ \mathbf{0}_{l \times 3n} & E \end{bmatrix}, \\ \bar{B}(\bar{z}) &= \begin{bmatrix} B(z) \\ \mathbf{0}_l \end{bmatrix}, \quad \bar{D}(\bar{z}) = \begin{bmatrix} D(z) \\ \mathbf{0}_{l \times 1} \end{bmatrix} \end{aligned} \quad (22)$$

where $\mathbf{0}_{3n \times l}$, $\mathbf{0}_{l \times 3n}$, $\mathbf{0}_l$, and $\mathbf{0}_{l \times 1}$ denote zero matrices with proper dimensions. The cost function (11) becomes

$$\Gamma = \int_t^\infty [\bar{z}^T(v)Q\bar{z}(v) + u^T(v)Ru(v)]dv \quad (23)$$

where

$$Q = \begin{bmatrix} Q_1 & \mathbf{0}_n & \mathbf{0}_n & -Q_1 F \\ \mathbf{0}_n & Q_2 & \mathbf{0}_n & \mathbf{0}_{n \times l} \\ \mathbf{0}_n & \mathbf{0}_n & Q_3 & \mathbf{0}_{n \times l} \\ -F^T Q_1 & \mathbf{0}_{l \times n} & \mathbf{0}_{l \times n} & F^T Q_1 F \end{bmatrix} \quad (24)$$

It is easy to verify that $Q \geq 0$.

To minimize the cost function (23) subject to the constraint (21) is non-trivial, because $\bar{A}(\bar{z})$, $\bar{B}(\bar{z})$, and $\bar{D}(\bar{z})$ are nonlinear and unknown due to the involvement of unknown human and robot dynamics. In this paper, this optimization problem is solved by employing an approximate dynamic programming method, which will be detailed in the following subsection.

B. Dynamic Programming

In this subsection, we provide an outline of the dynamic programming method developed in [26] to solve the optimization problem discussed in the previous subsection. Readers who are interested in the development of this method in its entirety may refer to [26]. Approximate dynamic programming mimics the way that biological systems interact with environments [27]. It usually includes an actor which generates an action according to the environment stimuli, and a critic that evaluates the result of the action.

In particular, the optimal control input u^* that minimizes the cost function (23) subject to the system constraint (21) is given by

$$u^*(\bar{z}) = -\frac{1}{2}R^{-1}\bar{B}^T(\bar{z})\Gamma_{\bar{z}}^* \quad (25)$$

where $\Gamma_{\bar{z}}^* = \frac{\partial \Gamma^*}{\partial \bar{z}}$ and $\Gamma^* = \min_{u(\bar{z})} \Gamma$ is the optimal cost function. However, as $\bar{B}^T(\bar{z})$ and Γ^* are unknown due to unknown dynamics of robot and human, the above optimal control needs to be obtained by approximation.

First, an observer has been designed to identify the system with unknown dynamics, as below

$$\dot{\hat{z}} = \bar{A}_0 \hat{z} + \hat{G}(\hat{z}, u) + L(\bar{z} - \hat{z}) \quad (26)$$

where \bar{A}_0 is a given Hurwitz matrix, $\hat{G}(\hat{z}, u)$ is the estimate of $G = \bar{A}(\bar{z})\bar{z} + \bar{B}(\bar{z})u + \bar{D}(\bar{z}) - \bar{A}_0 \hat{z}$, \hat{z} is the output of the observer, and L is chosen such that $A - L$ is also a Hurwitz matrix. $\hat{G}(\hat{z}, u)$ is approximated using a three-layer neural networks (NN) with a hidden layer, as below

$$\begin{aligned} \hat{G}(r) &= \hat{W}S(\hat{V}r), \\ S(\hat{V}r) &= [s_1(\hat{V}r), s_2(\hat{V}r), \dots, s_p(\hat{V}r)]^T, \\ s_i(\hat{V}r) &= \exp\left[-\frac{(\hat{V}r - \mu_i)^T(\hat{V}r - \mu_i)}{\eta_i^2}\right], \\ i &= 1, 2, \dots, p \end{aligned} \quad (27)$$

where $r = [\hat{z}^T u^T]^T$ is the NN input, p is nodes number of the hidden layer, $\mu_i = [\mu_{i,1}, \mu_{i,2}, \dots, \mu_{i,k}]^T$ is the center of

the receptive field, η_i is the width of the Gaussian function, and $\hat{W} \in \mathbb{R}^{(3n+l) \times p}$ and $\hat{V} \in \mathbb{R}^{k \times (4n+l)}$ are estimates of ideal weights W and V , respectively. \hat{W} and \hat{V} are updated online according to the following learning law

$$\begin{aligned} \dot{\hat{W}} &= -k_1(\bar{A}_0 - L)^{-T}(\bar{z} - \hat{z})S^T(\hat{V}r) - k_2\|\bar{z} - \hat{z}\|\hat{W} \\ \dot{\hat{V}} &= -k_3(I_p - \text{diag}(s_i^2(\hat{V}r)))^T \hat{W}^T(\bar{A}_0 - L)^{-T} \times \\ &\quad (\bar{z} - \hat{z})\text{sgn}^T(r) - k_4\|\bar{z} - \hat{z}\|\hat{V} \end{aligned} \quad (28)$$

where k_i , $i = 1, 2, 3, 4$ are positive scalars, and $\text{sgn}(r)$ denotes the sign function applied to all the components of r .

With the above approximation of the system dynamics, the control input u , which is the estimate of the optimal control u^* , can be obtained as

$$u = -\frac{1}{2}R^{-1}\left(\frac{\partial \hat{G}(\hat{z}, u)}{\partial u}\right)^T \hat{\Gamma}_{\hat{z}} \quad (29)$$

where $\frac{\partial \hat{G}(\hat{z}, u)}{\partial u}$ is obtained based on the above observer, as below

$$\begin{aligned} \frac{\partial \hat{G}(\hat{z}, u)}{\partial u} &= \hat{W} \frac{\partial S(\hat{V}r)}{\partial u} \\ &= \hat{W} \frac{\partial S(\hat{V}r)}{\partial(\hat{V}r)} \frac{\partial(\hat{V}r)}{\partial u} \\ &= \hat{W} \frac{\partial S(\hat{V}r)}{\partial(\hat{V}r)} \hat{V} \frac{\partial r}{\partial u} \\ &= \hat{W} \left[-\frac{2}{\eta_1^2}(\hat{V}r - \mu_1)s_1 \dots \right. \\ &\quad \left. -\frac{2}{\eta_p^2}(\hat{V}r - \mu_p)s_p \right] \hat{V} \begin{bmatrix} \mathbf{0}_{(3n+l) \times n} \\ I_n \end{bmatrix} \end{aligned} \quad (30)$$

Second, another NN is used to approximate $\Gamma(\hat{z})$, as below

$$\hat{\Gamma}(\hat{z}) = \hat{W}_1^T S_1(V_1^T \hat{z}) \quad (31)$$

where \hat{W}_1 is the estimate of the ideal weight W_1 , V_1 is a properly selected matrix, and $S_1(V_1^T \hat{z})$ is the activation function defined similar to $S(\hat{V}r)$ in Equ. (27). \hat{W}_1 is updated online according to the following learning law

$$\dot{\hat{W}}_1 = -k_5 \frac{\xi}{(\xi^T \xi + 1)^2} (\xi^T \hat{W}_1 + \hat{z}^T Q \hat{z} + u^T R u) \quad (32)$$

where k_5 is a positive scalar and

$$\begin{aligned} \xi &= \left[\frac{\partial S_1^T(V_1^T \hat{z})}{\partial(\hat{z})} \right]^T \hat{z} \\ &= \left[-\frac{2}{\eta_{1,1}^2}(V_1^T \hat{z} - \mu_{1,1})s_{1,1} \dots \right. \\ &\quad \left. -\frac{2}{\eta_{1,p_1}^2}(V_1^T \hat{z} - \mu_{1,p_1})s_{p_1} \right]^T V_1^T \hat{z} \end{aligned} \quad (33)$$

where $\eta_{1,j}$, $\mu_{1,j}$, $j = 1, 2, \dots, p_1$, and p_1 are NN parameters that are similar as in Equ. (27). Thus, $\hat{\Gamma}_{\hat{z}}$ is obtained as

$$\hat{\Gamma}_{\hat{z}} = \frac{\partial \hat{\Gamma}(\hat{z})}{\partial \hat{z}} = \frac{\partial S_1^T(V_1^T \hat{z})}{\partial \hat{z}} \hat{W}_1 \quad (34)$$

In short, the control input (29) with the observer (26) and updating laws (28) and (32) minimizes the cost function (23)

subject to the constraint (21) [26]. Therefore, the control objective discussed in Section II-B is achieved. Furthermore, roles of human and robot can be easily adjusted by changing open parameters Q_1 , Q_2 , Q_3 and R .

IV. SIMULATION STUDY

The simulation scenario is sketched in Fig. 1, where a human hand holds the end-effector of a planar robot arm with two revolute joints. The robot arm has a prescribed desired trajectory. There is a possible situation that an unexpected obstacle appears in halfway of the planned desired trajectory of the robot arm. Therefore, human needs to move the robot arm to equilibrium positions of his/her arm to avoid the obstacle. During this process, human and robot compete/cooperate with each other and the actual trajectory will be a compromise of the desired trajectory of the robot arm and the equilibrium positions of the human arm.

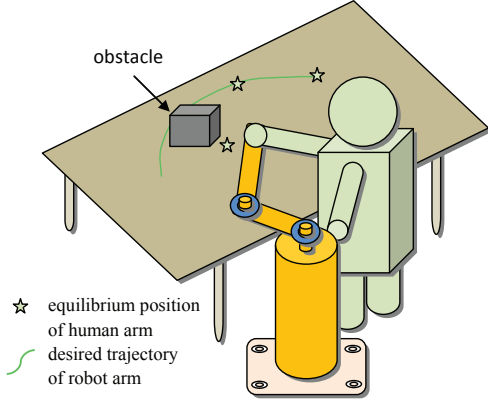


Fig. 1. Simulation scenario

A. Parameters

The parameters of the robot arm are given in Table I, where m_i , l_i , and I_{zi} , $i = 1, 2$, represent the mass, the length, and the inertia about the z-axis that comes out of the page passing through the center of mass, respectively.

TABLE I
PARAMETERS OF ROBOT ARM

Parameter	Description	Value
m_1	Mass of link 1	20.00kg
m_2	Mass of link 2	20.00kg
l_1	Length of link 1	0.30m
l_2	Length of link 2	0.30m
I_{z1}	Inertia of link 1	0.45kgm ²
I_{z2}	Inertia of link 2	0.45kgm ²

The impedance parameters of the human arm C_H and K_H are assumed to be diagonal, and their values are determined from [28]: $(21 \pm 20)\text{Ns/m}$ and $(201 \pm 200)\text{N/m}$, respectively. According to Property 1, we set

$$\begin{aligned} C_H &= \text{diag}\{[21 - 20 \sin(\dot{x}_1), 21 - 20 \sin(\dot{x}_2)]\} \\ K_H &= \text{diag}\{[201 - 200 \sin(\dot{x}_1), 201 - 200 \sin(\dot{x}_2)]\} \end{aligned} \quad (35)$$

where \dot{x}_1 and \dot{x}_2 are two elements of the velocity \dot{x} . The human intervention period is from $\frac{100}{3}\text{s}$ to $\frac{200}{3}\text{s}$, i.e., $C_H = K_H = \mathbf{0}_2$ for $t < \frac{100}{3}\text{s}$ and $t > \frac{200}{3}\text{s}$.

The desired trajectory of the robot arm is half a circle with a radius of 0.2m, from $[-0.2 \ 0]^T$ ($t = 0\text{s}$) to $[0.2 \ 0]^T$ ($t = 100\text{s}$). In particular, $x_d = [-0.2 \cos(\frac{\pi}{100}t) \ 0.2 \sin(\frac{\pi}{100}t)]^T$ is generated by Equ. (20) with

$$E = \begin{bmatrix} 0 & \frac{\pi}{100} \\ -\frac{\pi}{100} & 0 \end{bmatrix}, \quad F = \frac{20}{\pi} I_2 \quad (36)$$

A virtual obstacle is considered to be a square with the following four edges: $[-0.05 \ 0.2]^T$, $[0.05 \ 0.2]^T$, $[0.05 \ 0.1]^T$, and $[-0.05 \ 0.1]^T$. In order to avoid this virtual obstacle, two equilibrium positions of the human arm are given by: $[0 \ 0]^T$ ($t = \frac{100}{3}\text{s}$) and $[0.1 \ 0.1\sqrt{3}]^T$ ($t = \frac{200}{3}\text{s}$). The initial position of the robot arm in the joint space is $[-\frac{\pi}{2} \ \pi]^T$.

In order to obtain different role allocations of human and robot, different sets of Q and R in Equ. (23) are selected. The criterion of role allocation is as follows: when the actual trajectory is near to the prescribed desired trajectory, the tracking precision is more important so the robot arm should lead the task; otherwise, the human arm leads. For the convenience of analysis, the work space \mathcal{WS} is divided into two subspaces: $\mathcal{WS}_1 = \{x \mid \|x - x_d\| < 0.05\}$ in which the robot arm leads, and $\mathcal{WS}_2 = \mathcal{WS} - \mathcal{WS}_1$ in which the human arm leads. Furthermore, $Q_1 = 100I_2$ when $x \in \mathcal{WS}_1$ and $Q_1 = I_2$ when $x \in \mathcal{WS}_2$, and $Q_2 = 0.1I_2$, $Q_3 = 0.01I_2$, and $R = 0.001I_2$ in both subspaces. This is referred to as the ‘‘adaptive’’ case. For the comparison purpose, we also consider another two cases with two fixed roles for $x \in \mathcal{WS}$: 1) ‘‘human leading’’ with $Q_1 = 100I_2$; and 2) ‘‘robot leading’’ with $Q_1 = I_2$. These simulation settings are illustrated in Fig. 2.

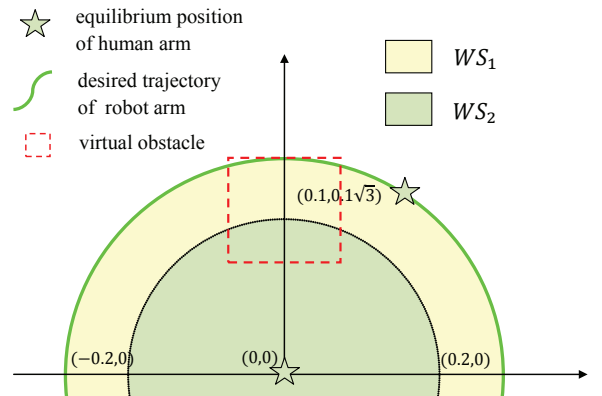


Fig. 2. Simulation settings

The Hurwitz matrix \bar{A}_0 in Equ. (26) is given as $\bar{A}_0 = \text{diag}[-I_6, E]$, $L = 10I_8$, and the initial values in Equ. (26) are: $u(0) = [0 \ 0]^T$, $\hat{z}(0) = [-0.2 \ 0 \ 0 \ 0 \ 0 \ 0 \ -\frac{\pi}{100} \ 0]^T$, $\hat{W}(0) = \mathbf{1}_{8 \times 10}$, $\hat{V}(0) = I_{10}$, $\hat{W}_1(0) = \mathbf{0}_{10 \times 1}$, and $V_1 = I_8$, where $\mathbf{1}_{8 \times 10}$ stands for a matrix with all elements as 1. The NN parameters are: $p = 10$, $p_1 = 10$, $\eta_i = 1$, $\mu_i = -1 + 0.2i$ for $i = 1, 2, \dots, p$, and $\eta_{1,j} = 1$, $\mu_{1,j} = -1 + 0.2j$ for $j = 1, 2, \dots, p_1$. The parameters in the updating laws (28)

and (32) are: $k_i = 1$ for $i = 1, \dots, 5$. A small sweeping frequency signal is added into the u to satisfy the condition of persistent excitation.

B. Results

The simulation results are shown in Figs. 3-6. From Figs. 3 and 4, it can be seen that different role allocations are realized in different task stages for the “adaptive” case, and the actual trajectory is a compromise of the desired trajectory of robot arm and equilibrium positions of human arm. Three stages are roughly divided by two role switches at $t = 37s$ and $t = 57s$, as illustrated in Fig. 4:

- In the beginning of the first stage, the human intervention is not applied till $t = \frac{100}{3}s$, so the actual trajectory tracks the desired trajectory of the robot arm after the adaptation of NN. Even when the human intervention is applied after $t = \frac{100}{3}s$, the robot arm still keeps its leading role because the actual trajectory is still within WS_1 .
- When human tries very hard (continuously applying a relatively larger force, as shown in Fig. 5) to move the robot arm from WS_1 to WS_2 , the leading role switches and the actual trajectory goes to the first equilibrium position of the human arm. The second equilibrium position is on the desired trajectory of the robot arm, so the robot arm moves to the second equilibrium position in a cooperative manner (generating a relatively smaller force).
- When the actual trajectory goes back into WS_1 , the leading role is switched back to the robot arm and the actual trajectory tracks the desired trajectory of the robot arm.

The above results are coherent with expectations. When the robot arm is near to the prescribed desired trajectory, Q_1 is given a relatively large value and the leader role is allocated to the robot arm. As a result, the actual trajectory of the robot arm (also the trajectory of the human arm) tracks the desired trajectory of the robot arm. When the human arm tries to move the robot arm to its equilibrium positions, a large interaction force will be generated, as shown in Fig. 5. Note that the property of impedance parameters of the human arm (stiffness and damping) is prescribed for the simulation purpose. In a practical situation, human may reduce his/her arm impedance to follow the robot arm if he/she finds that the desired trajectory of the robot arm is acceptable. When the robot arm is far away from the prescribed desired trajectory, Q_1 is given a relatively small value and the weight of the interaction force (i.e., Q_3) becomes relatively large. In this case, the human arm leads the task and the robot arm is moved to equilibrium positions with a small interaction force.

To demonstrate the significance of the proposed method, results for fixed roles, i.e., “human leading” and “robot leading”, are also shown in Figs. 3, 5, and 6. In the “human leading” case, a relatively smaller interaction force is found in Fig. 5. This indicates a good performance in the sense of human effort minimization. However, it also leads to a relatively larger tracking error even when there

is no human intervention, which is found in Figs. 3 and 6. Conversely, a relatively larger interaction force and a smaller tracking error are obtained in the “robot leading” case. These results indicate that a fixed role, i.e., either “human leading” or “robot leading”, cannot achieve the following expected performance: a smaller tracking error when there is no human intervention and a smaller interaction force when there is human intervention. It can be only achieved with an adaptive role, as shown in Figs. 3, 5, and 6.

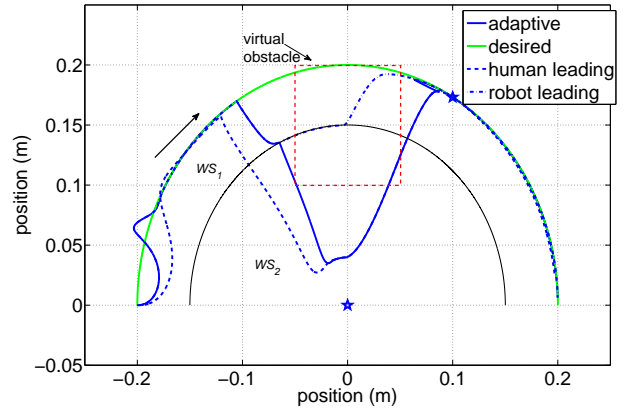


Fig. 3. Actual trajectories for 1) “adaptive” case, 2) “human leading” case, and 3) “robot leading” case, and desired trajectory of robot arm and equilibrium positions of human arm. The arrow denotes the motion direction of the end-effector of the robot arm.

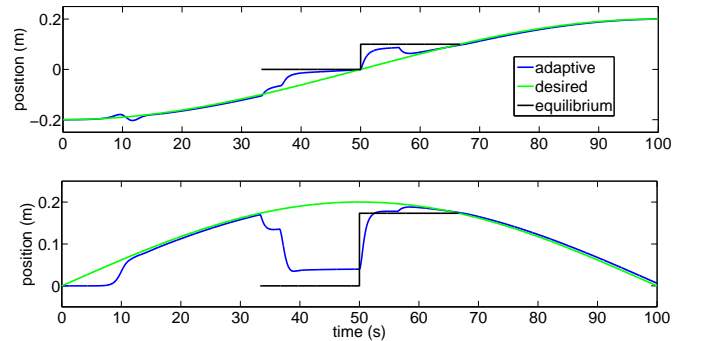


Fig. 4. Actual trajectory for “adaptive” case, desired trajectory of robot arm, and equilibrium positions of human arm in X direction (upper) and Y direction (below)

To intuitively understand the performances for cases of fixed and adaptive roles, we show three measures in Fig. 7: $\sum_0^{t_f} \|e\|$ where $e = x - x_d$, $\sum_0^{t_f} \|f\|$, and a measure of obstacle avoidance which is the sum of instances when the end-effector falls in the interior of the virtual obstacle. The performance is better if any of the three measures is smaller. It is clearly found that a worst performance will be resulted in if a best performance is achieved in the fixed role case, while the overall best performance is achieved in the adaptive role case.

C. Limitations

In above simulation studies, the bi-case switching of Q_1 is based on the distance of the robot arm to the prescribed

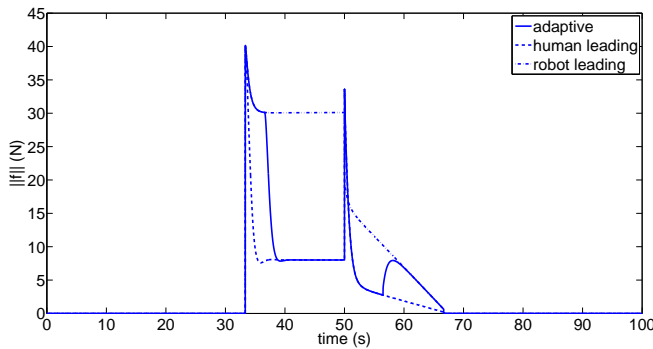


Fig. 5. Norms of interaction forces for 1) “adaptive” case, 2) “human leading” case, and 3) “robot leading” case

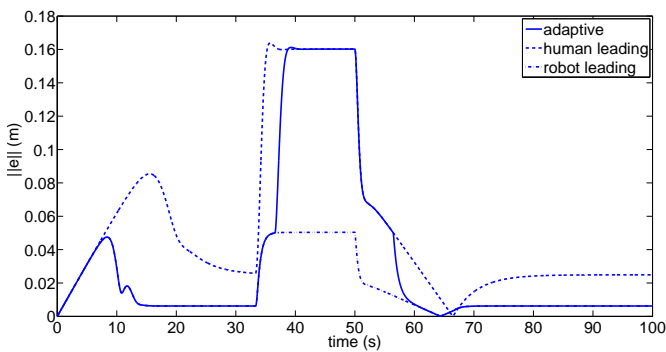


Fig. 6. Norms of tracking errors for 1) “adaptive” case, 2) “human leading” case, and 3) “robot leading” case

desired trajectory, which is heuristic and results in a non-smooth motion during the phase transition. Further effort will be made on determining the values of Q and R in different situations automatically, which is task-dependent and non-trivial. Besides, the human dynamics have been taken into account in the proposed method, which makes the investigation of concurrent adaptation become possible. However, real human adaptation may have not been adequately described in simulation studies, so future works of real-world experiments need to be conducted.

V. CONCLUSIONS

A framework of human-robot shared control has been developed in this work. The objectives of both agents have been taken into account in the problem formulation. An approximate dynamic programming method has been employed to provide an optimal control subject to unknown system dynamics. The validity of the proposed method has been verified by simulation studies.

REFERENCES

- [1] N. Hogan, “Impedance control: an approach to manipulation-Part I: Theory; Part II: Implementation; Part III: Applications,” *Transaction ASME J. Dynamic Systems, Measurement and Control*, vol. 107, no. 1, pp. 1–24, 1985.
- [2] M. M. Rahman, R. Ikeura, and K. Mizutani, “Investigation of the impedance characteristic of human arm for development of robots to cooperate with humans,” *JSME International Journal Series C*, vol. 45, no. 2, pp. 510–518, 2002.
- [3] V. Duchaine and C. Gosselin, “Safe, stable and intuitive control for physical human-robot interaction,” in *IEEE International Conference on Robotics and Automation*, pp. 3676–3681, 2009.
- [4] N. Jarrassé, V. Sanguinetti, and E. Burdet, “Slaves no longer: review on role assignment for human-robot joint motor action,” *Adaptive Behavior*, vol. 22, no. 1, pp. 70–82, 2014.
- [5] B. Corteville, E. Aertbelien, H. Bruyninckx, J. De Schutter, and H. Van Brussel, “Human-inspired robot assistant for fast point-to-point movements,” in *IEEE International Conference on Robotics and Automation*, (Roma, Italy), pp. 3639–3644, 2007.
- [6] C. Passenberg, N. Stefanov, A. Peer, and M. Buss, “Enhancing task classification in human-machine collaborative teleoperation systems by real-time evaluation of an agreement criterion,” in *IEEE World Haptics Conference*, pp. 493–498, 2011.
- [7] J. R. Medina, M. Lawitzky, A. Mortl, D. Lee, and S. Hirche, “An experience-driven robotic assistant acquiring human knowledge to improve haptic cooperation,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2416–2422, 2011.
- [8] M. S. Erden and T. Tomiyama, “Human-intent detection and physically interactive control of a robot without force sensors,” *IEEE Transactions on Robotics*, vol. 26, no. 2, pp. 370–382, 2010.
- [9] K. Wakita, J. Huang, P. Di, K. Sekiyama, and T. Fukuda, “Human-walking-intention-based motion control of an omnidirectional-type cane robot,” *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 1, pp. 285–296, 2013.
- [10] Y. Li and S. S. Ge, “Human-robot collaboration based on motion intention estimation,” *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007–1014, 2014.
- [11] J. R. Medina, M. Lawitzky, A. Molin, and S. Hirche, “Dynamic strategy selection for physical robotic assistance in partially known tasks,” in *IEEE International Conference on Robotics and Automation*, pp. 1180–1186, 2013.
- [12] R. Chipalkatty and M. Egerstedt, “Human-in-the-loop: Terminal constraint receding horizon control with human inputs,” in *IEEE International Conference on Robotics and Automation*, pp. 2712–2717, 2010.
- [13] R. Chipalkatty, G. Droge, and M. B. Egerstedt, “Less is more: Mixed-initiative model-predictive control with human inputs,” *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 695–703, 2013.
- [14] P. Evrard and A. Kheddar, “Homotopy switching model for dyadic haptic interaction in physical collaborative tasks,” in *3rd Joint Euro-Haptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (Salt Lake City, Utah, USA), pp. 45–50, 2009.
- [15] N. Jarrassé, T. Charalambous, and E. Burdet, “A framework to describe, analyze and generate interactive motor behaviors,” *PLoS ONE*, vol. 7, no. 11, p. e49945, 2013.
- [16] F. L. Lewis and D. Vrabie, “Reinforcement learning and adaptive dynamic programming for feedback control,” *IEEE Circuits and Systems Magazine*, vol. 9, no. 3, pp. 32–50, 2009.
- [17] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, “Reinforcement learning and feedback control: using natural decision methods to design optimal adaptive controllers,” *IEEE Circuits and Systems Magazine*, vol. 32, no. 6, pp. 76–105, 2012.
- [18] Y. Hirata, Z. Wang, K. Fukaya, and K. Kosuge, “Transporting an object by a passive mobile robot with servo brakes in cooperation with a human,” *Advanced Robotics*, vol. 23, pp. 387–404, 2009.
- [19] A. G. Feldman, “Functional tuning of the nervous system with control of movement or maintenance of a steady posture, ii: Controllable parameters of the muscles,” *Biophysics*, vol. 11, pp. 565–578, 1966.
- [20] F. A. Mussa-Ivaldi, N. Hogan, and E. Bizzi, “Neural, mechanical, and geometric factors subserving arm posture in humans,” *The Journal of Neuroscience*, vol. 5, no. 10, pp. 2732–2743, 1985.
- [21] T. Tsuji, P. G. Morasso, K. Goto, and K. Ito, “Human hand impedance characteristics during maintained posture,” *Biological Cybernetics*, vol. 72, no. 6, pp. 475–485, 1995.
- [22] K. P. Tee, E. Burdet, C. M. Chew, and T. E. Milner, “A model of force and impedance in human arm movements,” *Biological Cybernetics*, vol. 90, p. 368375, 2004.
- [23] K. P. Tee, D. W. Franklin, M. Kawato, T. E. Milner, and E. Burdet, “Concurrent adaptation of force and impedance in the redundant muscle system,” *Biological Cybernetics*, vol. 102, p. 3144, 2010.
- [24] C. Yang, G. Ganesh, S. Haddadin, S. Parusel, A. Albu-Schaeffer, and E. Burdet, “Human-like adaptation of force and impedance in stable and unstable interactions,” *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 918–930, 2011.

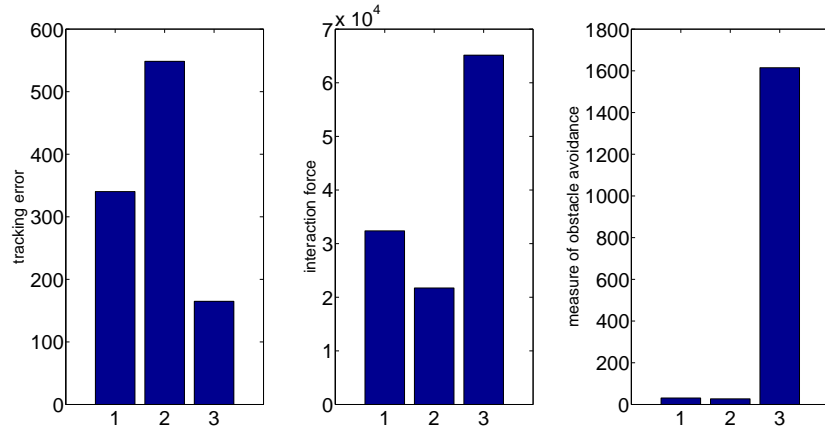


Fig. 7. Performance measures for 1) “adaptive” case, 2) “human leading” case, and 3) “robot leading” case

- [25] T. Tsumugiwa, R. Yokogawa, and K. Hara, “Variable impedance control based on estimation of human arm stiffness for human-robot cooperative calligraphic task,” in *IEEE International Conference on Robotics and Automation*, (Washington, DC, USA), pp. 644–650, 2002.
- [26] D. Liu, Y. Huang, D. Wang, and Q. Wei, “Neural-network-observer-based optimal control for unknown nonlinear systems using adaptive dynamic programming,” *International Journal of Control*, vol. 86, no. 9, pp. 1554–1566, 2013.
- [27] P. J. Werbos, “Intelligence in the brain: A theory of how it works and how to build it,” *Neural Networks*, vol. 22, no. 3, pp. 200–212, 2009.
- [28] S. P. Buerger and N. Hogan, “Complementary stability and loop shaping for improved human-robot interaction,” *IEEE Transactions on Robotics*, vol. 23, no. 2, pp. 232–244, 2007.